The 2nd Visual Object Tracking Segmentation VOTS2024 Challenge Results

Matej Kristan, Jiri Matas, Pavel Tokmakov, Michael Felsberg, Luka Čehovin Zajc, Alan Lukežič, Khanh-Tung Tran, Xuan-Son Vu, Johanna Bjorklund, Hyung Jin Chang, Gustavo Fernandez

University of Ljubljana
Faculty of Computer and Information Science

TOYOTA RESEARCH INSTITUTE

LINKÖPING UNIVERSITY

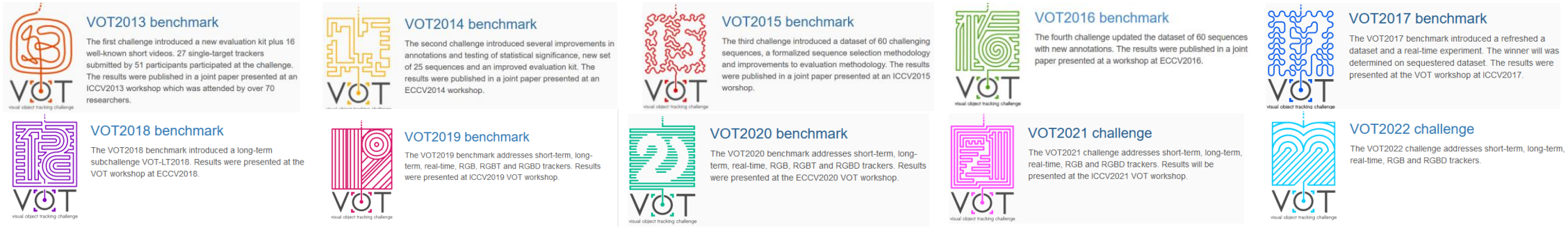
AUSTRIAN INSTITUTE OF TECHNOLOGY

UNIVERSITY OF BIRMINGHAM

VOTS2024 challenge

- The VOT formed in 2013 to support general object trackers
- Tracking objects that do not change their topology

  (don't break apart or turn inside-out)

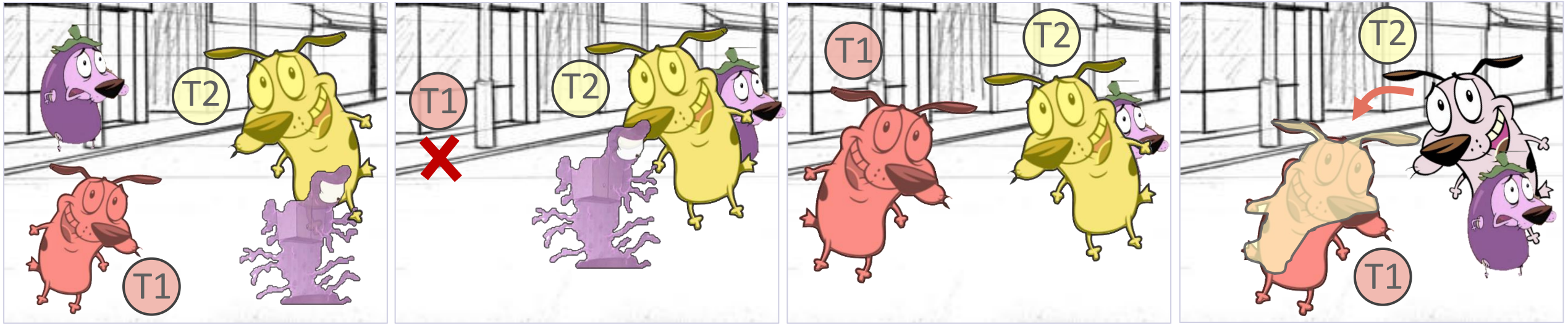


- In 2023 VOT challenges converged into VOTS:

  segmentation-based general object tracking no longer distinguishing between single-target, multi-target, short-term and long-term tracking

The VOTS2024 challenge scope

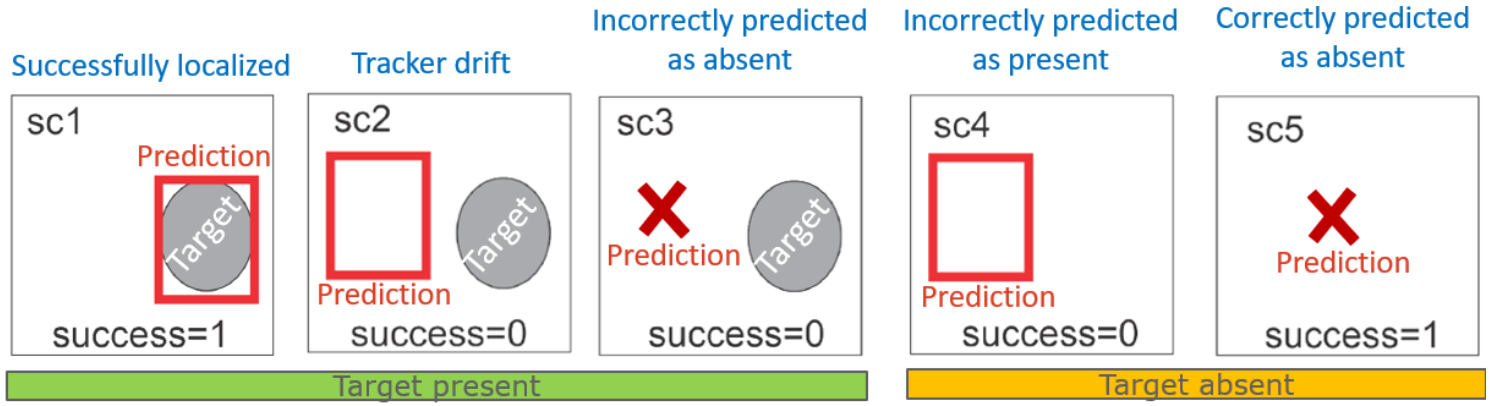
- General object Short/Long-term, Single/Multi-target segmentation trackers
- Initialize on all targets in the first frame and report position in the rest



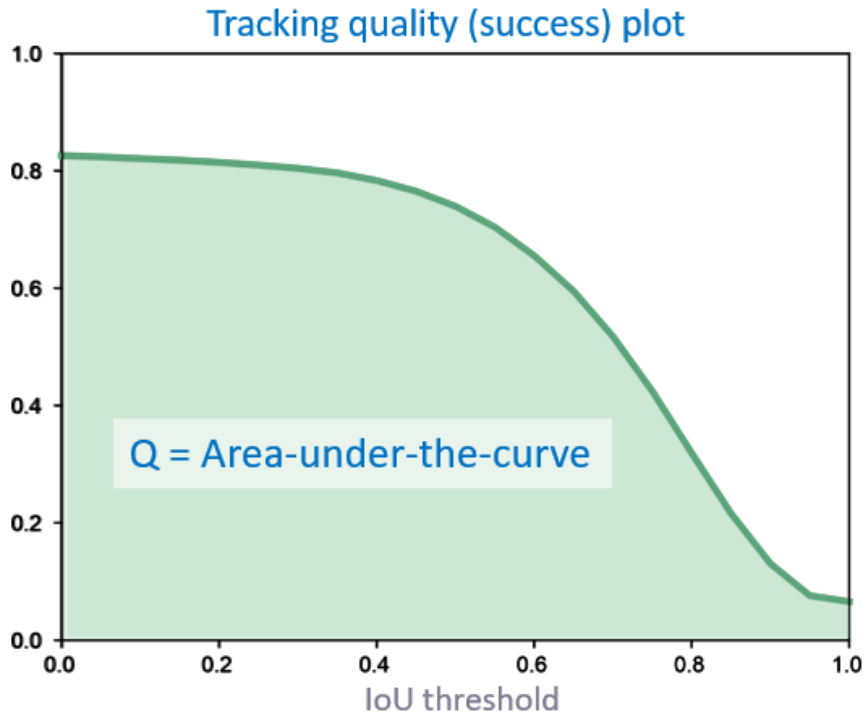
- Determine the target absence and redetect when it reappears
- Drifting off the target to background or another object is considered failure

VOTS Primary performance measure *Q*

- VOTS2023 *tracking success* accounts for 5 scenarios:



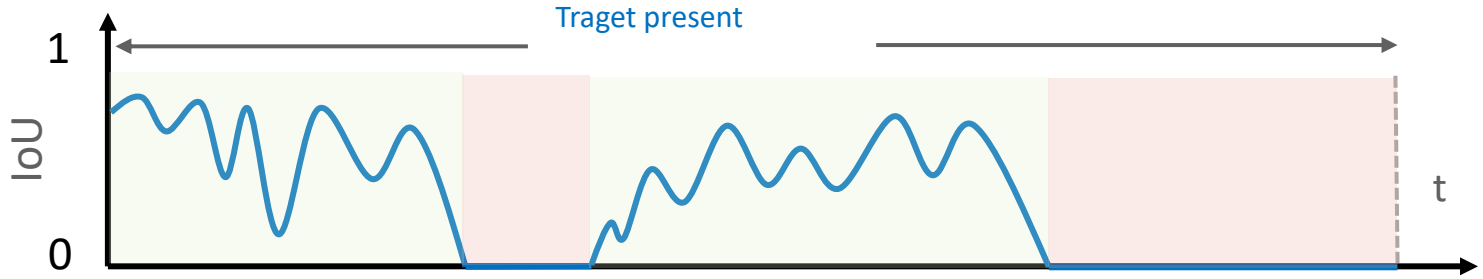
- Tracking quality plot:
  - Average of success plots calculated for each target in each sequence

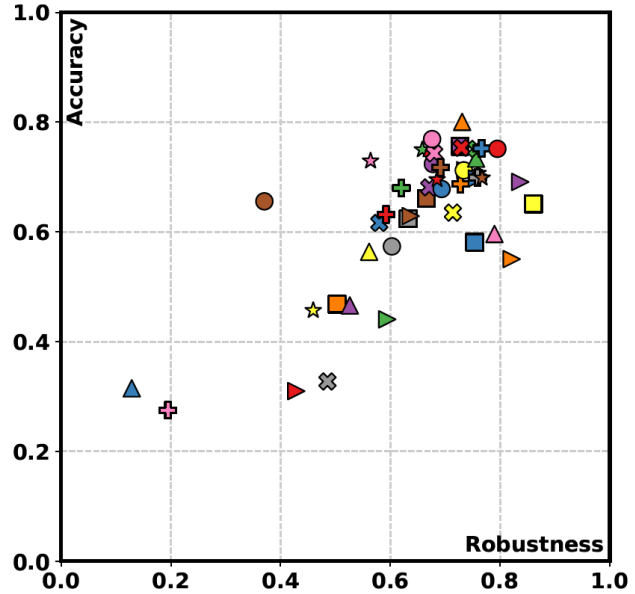


- Primary measure: *Tracking quality Q* (area-under-the-curve)

VOTS Auxiliary performance measures

- Accuracy/Robustness (@IoU=0.0 when target present)





"Why did the tracker fail while target visible?"

- NotReportedError (NRE): % frames incorrectly predicted target absent
- DriftRateError (DRE): % frames tracker drifted while predicting target present

"How well is target absence determined?"

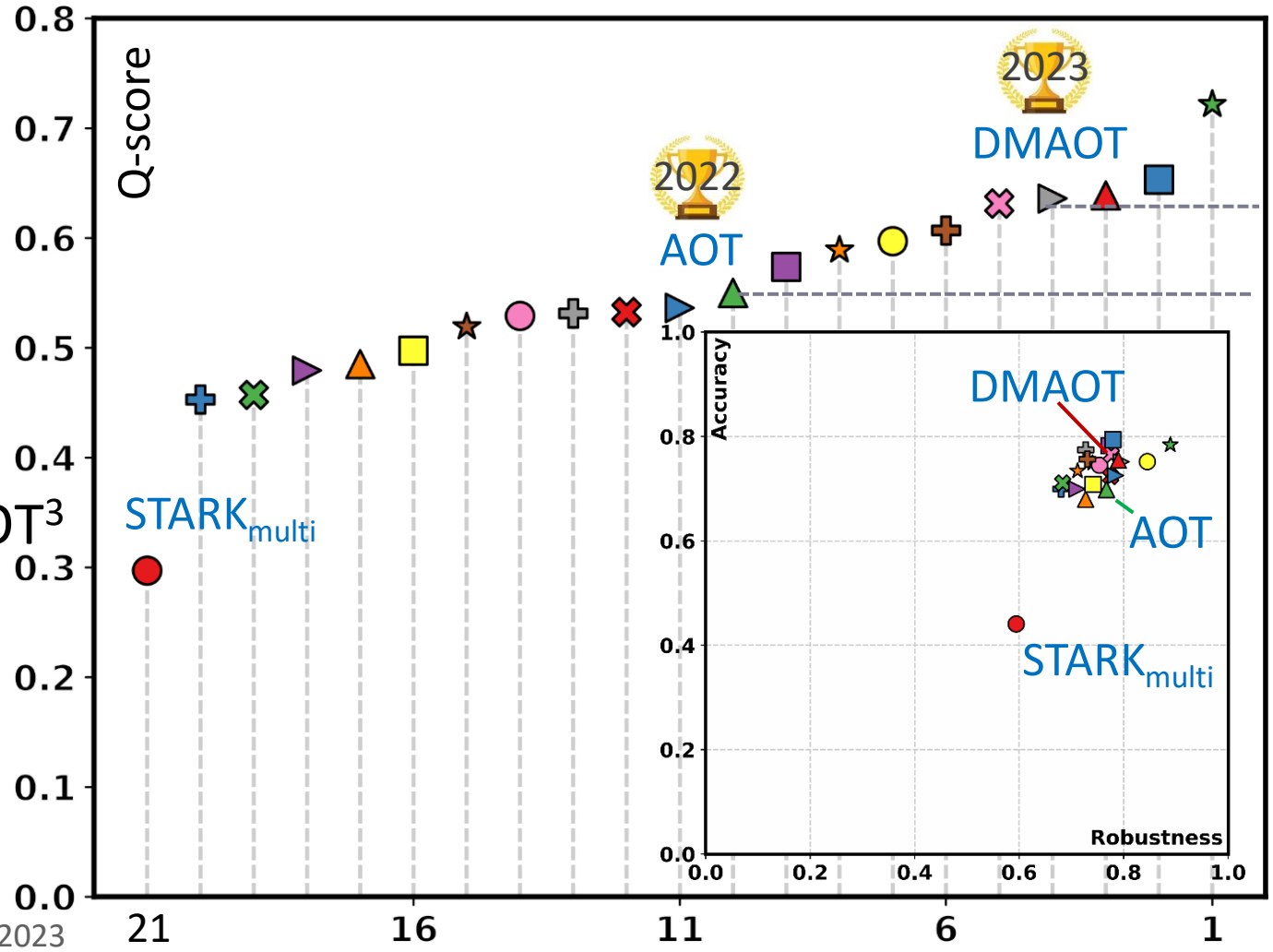
- AbsenceDetectionQuality (ADQ): % frames target correctly predicted absent

VOTS2024 dataset: same as VOTS2023

- Stats: 144 sequences ; 341 targets ; 168 targets leave the FOV at least once
- Sequence properties:
  - min/max = 63/10.7k frames
  - On average 2.37 targets per sequence annotated
  - Median target absence: 18 frames
- To prevent overfitting:
  - Sequences + initialization frames GT publicly available.
  - GT of test frames sequestered, evaluation carried out on a dedicated server.

VOTS2024 challenge results: 18 trackers tested

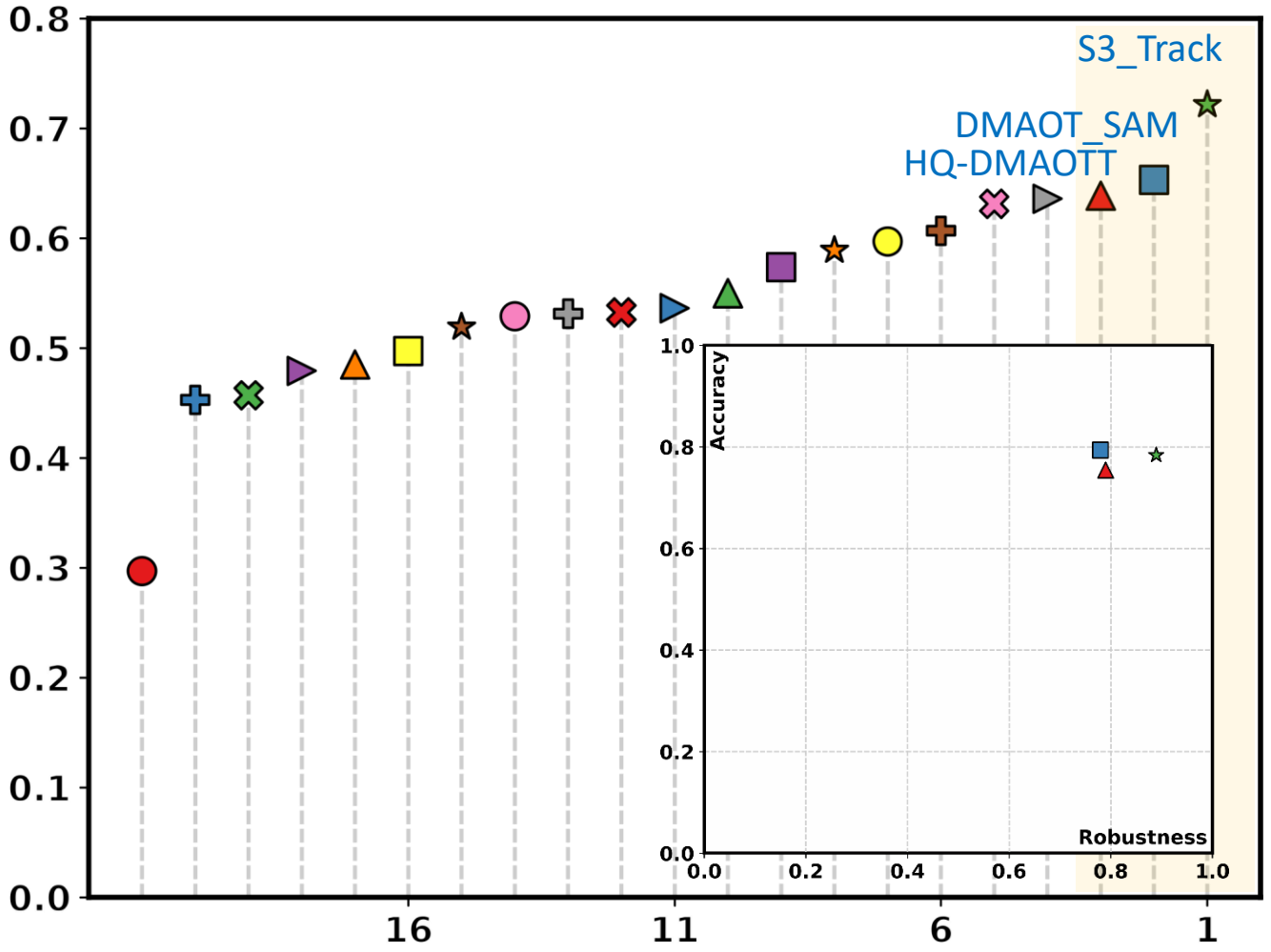
- Baseline 1: Independent STARKs[1] (41% in Q-score w.r.t. top tracker)
  all submissions outperform it
- Baseline 2: VOT2022 winner AOT[2]
  44% (8 trackers) outperfrom it
- Baseline 3: VOTS2023 winner DMAOT[3]
  17% (3 trackers) outperfrom it



[1] Yan, et al. ICCV2021; [2] Yang, et al. NeurIPS 2021 ; [3] Cheng, et al. VOTS 2023

VOTS2024 challenge top 3 trackers

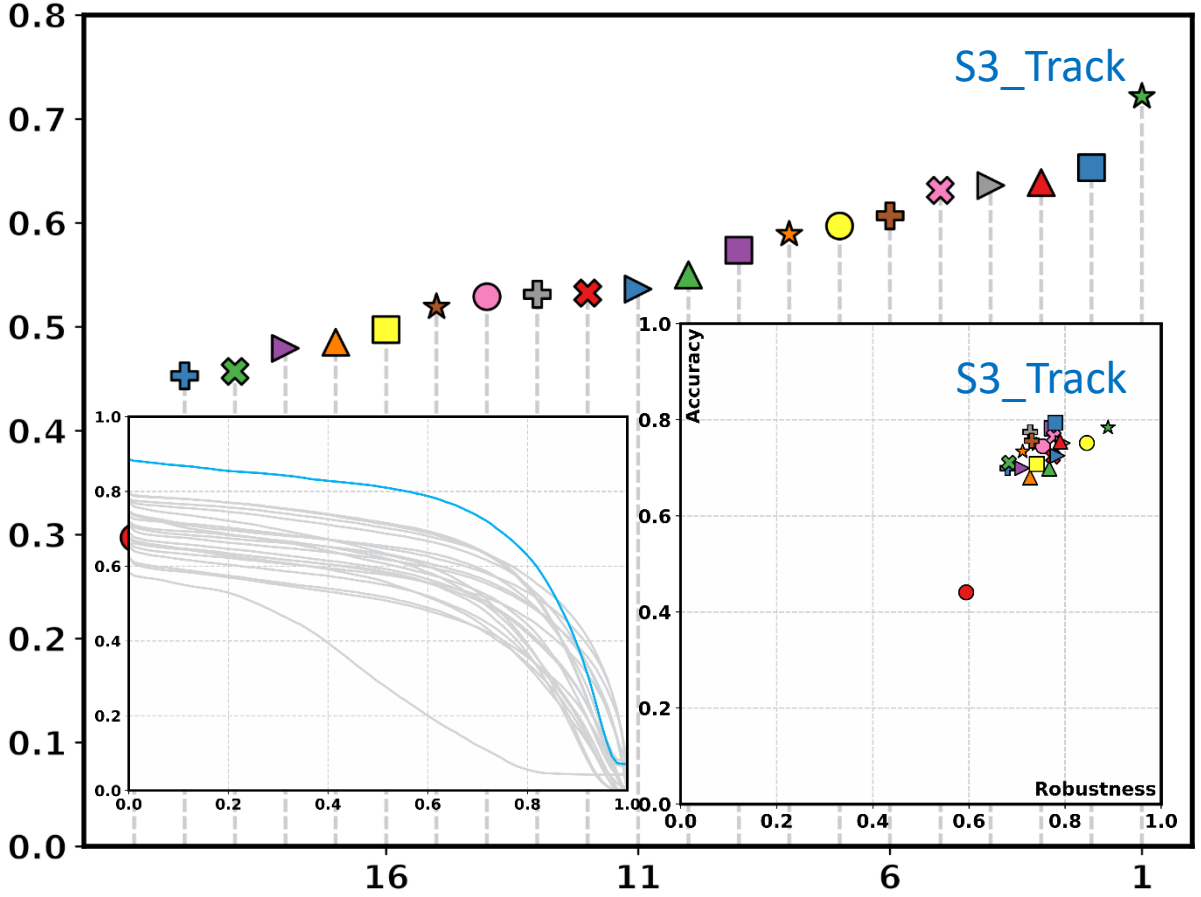
- Top 3 trackers: S3_Track, DMAOT_SAM, HQ-DMAOTT
  - S3_Track:
    - Single-stage architecture
    - Extends Cutie[1] by multi-scale features + ViT backbone +…
  - DMAOT_SAM & HQ-DMAOTT:
    - Two-stage architecture
    - DMAOT[1] for the initial mask prediction, then refine by SAM[3]
- All apply: ViT/Swin backbones + large training datasets



[1] Cheng, et al. CVPR2024; [2] Cheng, et al. VOTS 2023; Kirillov et. al. ICCV2023

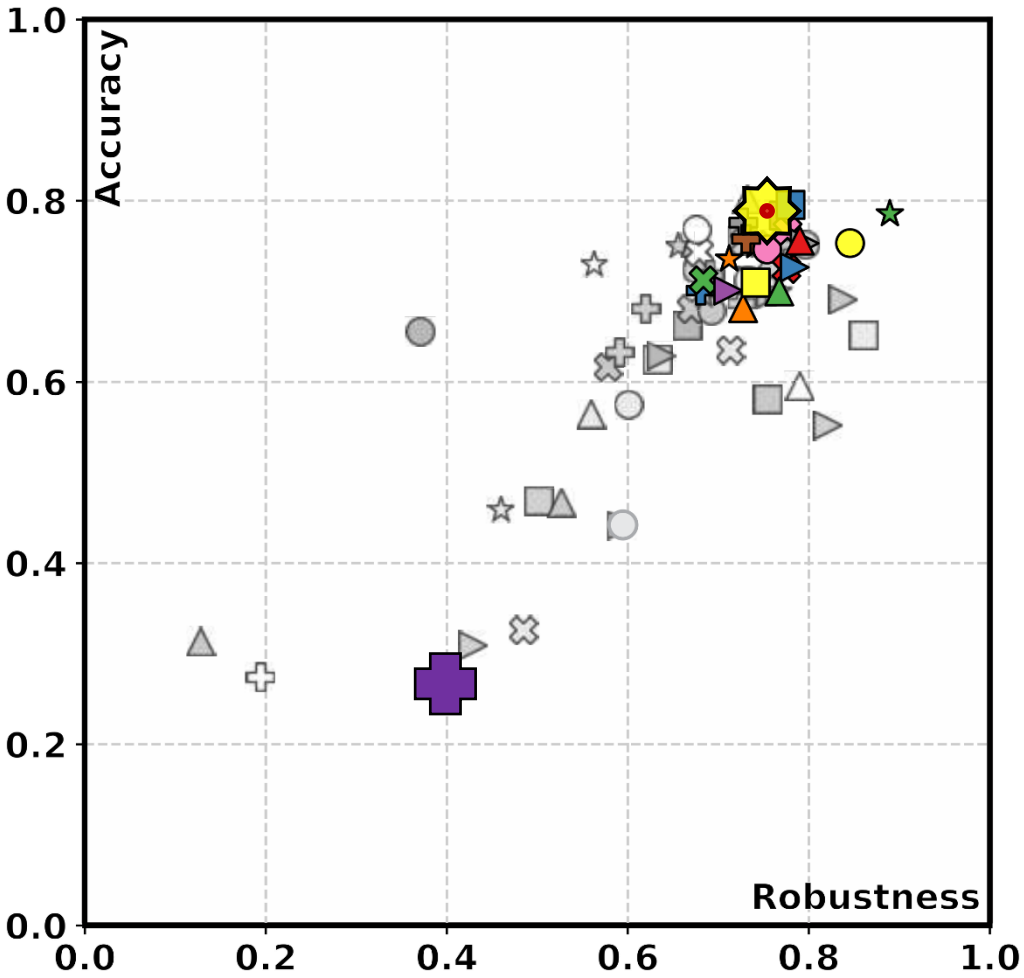
VOTS2024 top-ranked tracker S3_Track

- *Q-score = 0.722* (>10% improvement over second-best)
- VOTS aux measures: *A=0.784, R=0.889 , DRE=0.04 , NRE=0.07, ADQ=0.781*
- Good segmentation accuracy 78.4%
  (DMAOT_SAM; A=79.4%)
- On-target in ~89% of frames
- Localization failure 11% due to:
  - Falsely reported target absent: 63%
  - Drifting: 37%
- Accurately predicts target absence in 78% cases



VOTS2024 vs VOTS2023

- Compared to 2023, 2024 submissions improved on accuracy as well as robustness
  - VOTS2023 winner Q-score: 0.636
  - VOTS2024 top Q-score: 0.722
    (13% winner performance boost)
- "Ancient" history (10 years):
  - 2014 sota KCF[1]: Q=0.107, A=0.268, R=0.403
- SAM2[2]: Q=0.642, A=0.789, R=0.75
  - Outperforms last-year's winner by 1% in Q
  - Would have ranked 3rd in Q-score on VOTS2024



[1] Henriques, et al. ECCV2012; [2] Ravi, et al. Arxiv2024

VOTS2024 challenge
Winners:

S3_Track by: Deshui Miao, Xingsen Huang, Xin Li, Dongmei Jiang, Ming-Hsuan Yang, Yaowei Wang

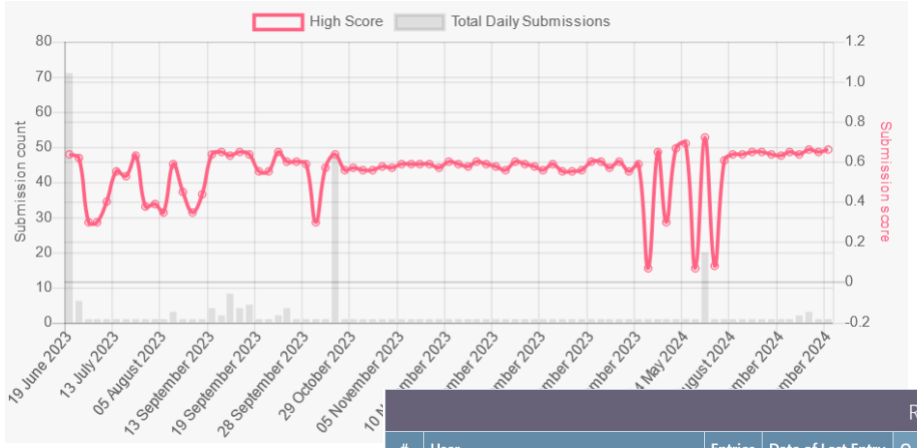
"Associating Semantic Tracker"

Winners talk end of Session I
@10:15

Summary

- Presented VOTS2024 results

- Evaluation server open for post-challenge evaluation



RESULTS

#	User	Entries	Date of Last Entry	Q ▲	Acc ▲	Rob ▲	NRE ▲	DRE ▲	ADQ ▲
1	VOTS2024_S3_Track	4	07/08/24	0.72 (1)	0.78 (5)	0.89 (1)	0.07 (36)	0.04 (74)	0.78 (6)
2	hello_world	11	09/22/24	0.66 (2)	0.79 (3)	0.79 (11)	0.14 (26)	0.07 (66)	0.72 (16)
3	civa_	1	09/20/24	0.66 (3)	0.78 (8)	0.80 (9)	0.14 (28)	0.06 (69)	0.76 (9)
4	civa_lab_	2	09/18/24	0.65 (4)	0.79 (4)	0.79 (12)	0.14 (25)	0.07 (65)	0.72 (15)
5	VOTS2024_dmaot_sam	1	07/08/24	0.65 (5)	0.79 (2)	0.78 (17)	0.14 (22)	0.08 (61)	0.73 (11)

VOTS2024 benchmark

Thanks

- The VOTS2024 committee


M. Kristan


J. Matas


P. Tokmakov


M. Felsberg


L. Čehovin Z.


A. Lukežič


T. Tran


Xuan-Son Vu


Johanna Björklund


H. J. Chang


G. Fernandez

- Everyone who participated or contributed

Minasadat Attari, Antoni Chan, Liang Chen, Xin Chen, Jaired Collins, Yutao Cui, Ganesh Sai Manas Devarapu, Yinglong Du, Heng Fan, Wan-Cyuan Fan, Zhenhua Feng, Mingqi Gao, Rama Krishna Sai Gorthi, Raghav Goyal, Jungong Han, Bijaya Hatuwal, Zhenyu He, Xiantao Hu, Xingsen Huang, Yuqing Huang, Dongmei Jiang, Ben Kang, Palaniappan Kannappan, Josef Kittler, Simiao Lai, Ning Li, Xiaohai Li, Xin Li, Cheng Liang, Liting Lin, Haibin Ling, Ting Liu, Ziquan Liu, Huchuan Lu, Yifei Luo, Deshui Miao, Juan Mogollon, Ziqi Pang, Jaswanth Reddy Pochimireddy, Viktor Prutyanov, Gani Rahmon, Aleksandr Romanov, Liangtao Shi, Mennatullah Siam, Leonid Sigal, Arun Kumar Sivapuram, Roman Solovyev, Elham Soltani Kazemi, Imad Eddine Toubal, Jia Wan, Limin Wang, Xinying Wang, Yaowei Wang, Yu-Xiong Wang, Zhiquan Wang, Gangshan Wu, Qiangqiang Wu, Xiaojun Wu, Zihao Xia, Jinxia Xie, Chenlong Xu, Tianyang Xu, Yong Xu, Chaocan Xue, Chao Yang, Jinyu Yang, Ming-Hsuan Yang, Chenyang Yu, Ke Yu, Chunhui Zhang, Jiaming Zhang, Zhipeng Zhang, Feng Zheng, Yaozong Zheng, Bineng Zhong, Jinglin Zhou, Junbao Zhou, Yong Zhou, Zikun Zhou, Guibo Zhu, Jiawen Zhu, Xuefeng Zhu, and Vladimir Zunin

VOTS2024 benchmark

- VOTS2024 sponsors:

University of Ljubljana Faculty of Computer and Information Science

arnes

WASP WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM

UNIVERSITY OF BIRMINGHAM